Excerpt from: DRAFT "Water Quality Standards Rule Packages Technical Support Document" for Biocriteria and Phosphorus Response Criteria Rules, excerpted on 7-21-2017, Wisconsin Dept. of Natural Resources

3.3 STREAM PHOSPHORUS RESPONSE CRITERIA

3.3.1 Datasets

The following datasets were used for development of stream phosphorus response criteria.

Nutrient Impacts (Version 2) Dataset. To determine which stream metrics have the strongest correlation to TP concentrations, and thus which would best represent the variables in the conceptual model, WDNR assembled existing data from three different studies spanning ten years. The 197 stream sites that were used for this analysis included 171 sites from the 2001-03 wadeable stream nutrient impacts study (Robertson et al. 2006), 8 sites from WDNR's 2007-09 watershed rotation study, and 18 sites from WDNR's 2011 high N:P ratio study. The sites were selected to span the range of nutrient conditions and to minimize the correlation between total phosphorus and total nitrogen. The dataset included a variety of metrics for fish, macroinvertebrates, and diatoms, and includes sites from each Natural Community and Ecoregion.

Reference Sites Dataset. To determine the natural range of healthy, or least impacted, streams, we used a reference site dataset. This dataset was collected during the 2008-2009 sampling seasons, with the express purpose of assembling a least impacted dataset for the state that provided statistically valid representation of all Natural Community Types across each Ecoregion. See the detailed description of this dataset in Section 3.1.2 of this document.

3.3.2 Benthic algal biomass & diatom taxa

For streams, primary productivity can be measured in one or both of the following ways. To maximize efficiency for making assessment decisions, the "viewing bucket" method for algal biomass is recommended as the first step in assessing primary productivity. If these results are conclusive, as described below, no further analysis is required. If the results are inconclusive (mid-range scores), further analysis of the diatom community is required to determine whether the stream is exhibiting a TP response.

a. Viewing Bucket for algal biomass

A visual assessment of benthic algal biomass in streams using a quantifiable system such as a viewing bucket is an efficient and appropriate screening tool to determine whether a site clearly is, or is not, exhibiting a nutrient response. High TP can be expected to result in greater biomass and coverage of benthic algae in streams. The viewing bucket method is included in the *U.S. EPA's Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers* (Barbour et al. 1999) and is used in several states' monitoring programs.

The method will be employed during evaluation of habitat assessment transects. Benthic algal biomass will be observed and characterized on a grid of 25 points with the viewing bucket (Figure 1). This will be done once on each of the ten habitat transects for a reach, staggered across the stream from left to right. Scores from each transect will then be averaged for the reach.

The assessment should be conducted during the growing season (July, Aug, or Sept) during baseflow conditions, with the first viewing bucket assessment in July or August, and second (if needed) in August or September. Because scouring during stream spate events may reduce algal biomass, sampling should be avoided within two weeks of a storm event.

Figure 1. An example viewing bucket from Rhode Island Dept. of Environmental Management. Photograph by A. Patterson.

Determination of criteria thresholds

Thresholds for algal biomass as evaluated with the viewing bucket method reflect the expectation that higher TP levels will lead to higher algal biomass. The viewing bucket scoring scale is from 0 (low biomass) to 3 (high biomass). If the average algal biomass score for the reach is less than 1, the stream is not impaired by TP and there is no need for further primary production assessment. If the algal biomass score is greater than 2, the stream is impaired by TP and no further assessment is necessary. If the algal biomass score is between 1 and 2, further primary production assessment via the Diatom Phosphorus Index (DPI) is needed.

Stream phosphorus response indicator for benthic algal biomass using viewing bucket method.

Benthic algal biomass, viewing bucket score (0-3)	Attainment decision	
	Aquatic Life Use	Recreation Use
< 1	Attained; no further primary production assessment	Attained
1 - 2	Inconclusive; assess benthic diatoms using DPI	
> 2	Not attained; no further assessment	Not attained

The viewing bucket method can also be used to assess whether a stream is attaining its recreation use, as recreation is also impacted by algal growth. A stream's recreation use is considered attained if the viewing bucket score is at 2 or below.

b. Diatom Phosphorus Index (DPI)

Diatoms are a form of algae with a silicate shell with many species that tend to be found on stream beds or clinging as a brown substance to filamentous algae, such as Cladophora. They are found in both

freshwater and marine waters and in many environments play a very substantial role in primary productivity within the system. Analysis of diatoms has been used for water quality analysis around the world. Various species have been identified as tolerant or sensitive to various stressors, including nutrients.

In development of phosphorus criteria for wadeable streams, WDNR used three indices to evaluate diatom community responses to phosphorus: the Diatom Nutrient Index (DNI), the Diatom Siltation Index (DSI), and the Diatom Biotic Index (DBI) (Robertson et al. 2006). Because these indices are primarily based on literature-derived tolerance values that are not specific to phosphorus, we decided to develop a new method that is specific to phosphorus and calibrated to Wisconsin diatom data, herein referred to as the Diatom Phosphorus Index (DPI).

The DPI is based on a statistical method called Weighted Averaging (WA; ter Braak and van Dam 1989). This method can be used to determine whether the diatom community at an assessment site resembles the community that is typically found at sites meeting the stream TP criterion. The TP criterion is based on breakpoints in the relationships between TP and diatom (and other biological) metrics, and as such represents the level of TP where the biological community changes the most.

WA estimates species-specific environmental preferences (optima) as the average value of an environmental variable (in this case, TP) where a species occurs, weighted by its relative abundance. The DPI at a site is then estimated as the weighted average of the TP optima of all the species present at that site. WA was developed to infer paleo-limnological characteristics such as pH, temperature, and TP (reviewed in Juggins and Birks 2012), and has also been used to develop a stream diatom nutrient index in New Jersey (Ponader 2007).

A WA model was developed from the Nutrient Impacts (Version 2) Dataset described above. Diatom and nutrient samples were collected in 2001-03 and 2011 using methods described in Robertson et al. (2006). Diatom samples were collected in September, and nutrient samples were collected monthly from May-Oct. Models using various subsets of nutrient samples during and prior to September were evaluated to determine whether they were better predictors of diatom community structure than the entire growing season, but the median of all six monthly samples was the best predictor. Only taxa with at least five occurrences (n=156) were used in the model development.

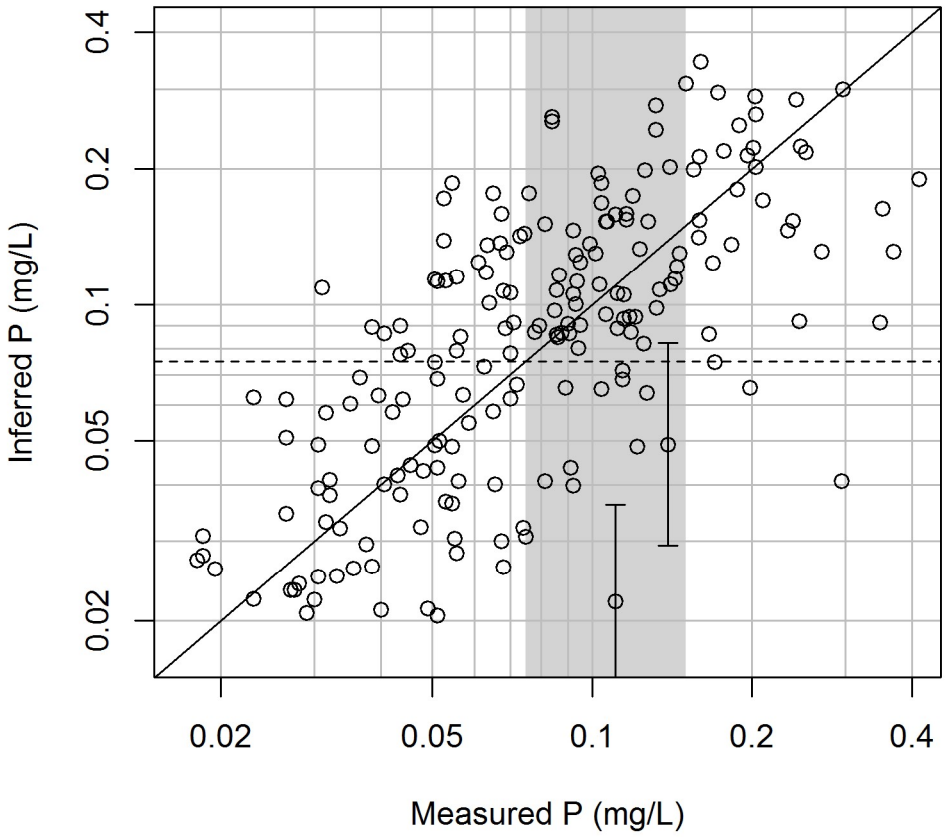
The WA model was fit using the WA function in the rioja package (Juggins 2014) in R. Prediction errors were estimated by leave-one-out cross-validation. The cross-validated r^2 is 0.49, which means that TP explains about half of the variation in diatom community structure among sites (Figure 2). The root mean square error of prediction (RMSEP) is 62%, which means that the average DPI differs from the measured TP by 62%. The residual variation in this relationship probably reflects sampling error in both TP and diatoms. The DPI may actually be a more accurate reflection of prevailing phosphorus conditions than direct stream TP measurements.

For purposes of assessing attainment of the diatom phosphorus response criteria, the department would not list a waterbody as impaired if it is 90% confident that the diatom community is not impaired. A bootstrapping procedure[1] was used to estimate confidence intervals around DPI values. If the upper 80% confidence limit of DPI is < 75 µg/L, we would be 90% confident that the diatom community is not impaired. Among the 68 sites in the model dataset where biological confirmation would be relevant (measured TP is 75-150 µg/L), only two would be considered not impaired through the perspective of the diatom community.

[1] In rioja, the predict function with sse=TRUE estimates standard errors for each site (v1), which reflect how much the inferred P varies across the bootstrapped samples.

For assessment purposes, the DPI should be used only in conjunction with TP, not as a stand-alone assessment metric. It has not been shown to be sensitive to a broader range of environmental stressors than phosphorus. However, high algal viewing bucket scores may be used to list a water as impaired regardless of TP concentrations.

Figure 2. Measured vs. diatom-inferred TP concentration from Weighted Average (WA) model (R2 = 0.49). Note log scales on both axes. Gray area is TP range where biological confirmation may be used. Error bars are 80% confidence intervals on two example points.



3.3.3 Other metrics not selected

For streams, we determined that primary production metrics were the most appropriate as phosphorus response criteria, and upper-level indicators did not add clear value to assessment of phosphorus response. We considered both macroinvertebrate and fish metrics, as described below, but determined that the relationships between these metrics and phosphorus, as assessed using currently available data, were not strong enough to include as response indicators.

The following primary productivity metrics were considered for development of stream phosphorus response criteria but were not selected.

- *Benthic* chlorophyll *a*
  Benthic chlorophyll *a* is a useful metric in streams; however, because it is not clear what an appropriate threshold value is, this metric is not proposed to be required by code at this time. Though it will not be required in code, staff are encouraged to collect benthic algae via a rock scrape

for chlorophyll *a* analysis, to be used as supplemental information and to help build WDNR's dataset on benthic chlorophyll *a*. Obtaining additional data will help the department refine benthic chlorophyll thresholds in the future.
    - When conducting a rock scrape, WDNR considered requiring a pebble count to determine what percent of the stream bed was available habitat for algae to grow on. However, for purposes of determining whether the system is exhibiting a response to TP, it was determined that even if there is only a small percentage of streambed with hard substrate but that substrate is showing significant algae growth, it should be considered as exhibiting a response to TP.
- *Sestonic* chlorophyll *a*
  Streams do not typically have high sestonic (suspended) chlorophyll *a* levels, so a grab sample of sestonic chlorophyll *a* is not needed for stream sites. Sestonic chlorophyll *a* is a more appropriate indicator for rivers.
- *Algal toxins*
  While production of algal toxins can be a result of high TP concentrations, algal toxins are not recommended as a primary phosphorus response indicator. High algal toxins are very rarely a problem in streams. At the current time, protocols for assessing algal toxins are insufficient. However, an algal toxin sample may be collected and analyzed in a stream if a problem is suspected, and the analysis may be used as supplementary evidence of a problem.
- *Diatom Nutrient Index (DNI) and Diatom Biotic Index (DBI)*
  The weighted average Diatom Phosphorus Index (DPI) was selected over the Diatom Nutrient Index (DNI) or Diatom Biotic Index (DBI) because it shows a stronger correlation with phosphorus.

Macroinvertebrate and Fish Metrics

To support the continued development of stream nutrient criteria, including phosphorus response criteria, WDNR conducted an extensive set of analyses on biotic responses to nutrients. This study is described in detail in a report titled "Evaluation of the relative effects of phosphorus and nitrogen on stream biological community structure" (Diebel 2015). One of the major conclusions of this study is that nitrogen and phosphorus have independent and statistically significant effects on the community structure of all taxonomic groups, but the strength of those effects is relatively weak compared to other environmental variables, except for the effect of P on diatoms, which is strong. In particular, both macroinvertebrate and fish communities are more strongly influenced by stream size, temperature, and conductivity than by phosphorus. These higher taxonomic groups are useful indicators of overall biological integrity in streams, but do not appear to be specific enough indicators of phosphorus to serve as response variables.